H

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/447,400	11/22/1999	DAVID T. CHEN	CE08051R	3769

7590 12/30/2003

JONATHAN P MEYER
MOTOROLA INC
CORPORATE OFFICES
1303 E ALGONQUIN ROAD
SCHAUMBURG, IL 60196

EXAMINER	
BAUGH, APRIL L	
ART UNIT	**PAPER NUMBER**
2141	10

DATE MAILED: 12/30/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/447,400	CHEN ET AL.	
	Examiner	Art Unit	
	April L Baugh	2141	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-14* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-14* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. §§ 119 and 120

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

13)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application) since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

a) ☐ The translation of the foreign language provisional application has been received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121 since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____.
4) ☐ Interview Summary (PTO-413) Paper No(s). _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other:

DETAILED ACTION

Response to Amendment

Claims 1-14 are pending.

Response to Arguments

1. Applicant's arguments with respect to claim1-4 in view of have been considered but are moot in view of the new ground(s) of rejection.

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claim1-4, 7, 8, and 10-14 rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent 6,078,943 to Yu in view of Dias et al. and further in view of Devarakonda et al.

Referring to claim 1, Yu teaches in a communication system network having a plurality of servers (Fig. 1, reference 54 of Yu), each of said plurality of servers having a load level based on serving a number of clients in said communication system network (column 1, lines 9-11 and 28-30 of Yu), a method comprising the steps of: grouping said plurality of servers into a first server group (column 4, lines 52-53 of Yu), wherein said first server group has a load level less than load level of said second server group (column 3, lines 53-55 and column 6, lines 55-58 of

Yu); calculating a time period T (column 3, line 1 of Yu); assigning load to a server selected from servers in said first server group from an initial time until expiration of said time period T (column 2, lines 25-27 and 58-60 of Yu).

Yu does not teach grouping said plurality of servers into a first and second server groups. Dias et al. teaches grouping said plurality of servers into a first and second server groups (column 2, lines 58-65 and column 3, lines 3-5 and column 4, lines 20-22). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu by grouping said plurality of servers into a first and second server groups because the creation of server groups based on server load allows request to be sent to a group where all servers within the group are just as capable of processing the request without overload thus creating request response time more efficient and maintaining load balance.

Yu in view of Dias et al. does not teach to assign load to a server after time T. Devarakonda et al. teaches assigning load to a server selected from servers in said first and second server groups after expiration of said time period T (column 2, lines 52-57). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu in view of Dias et al. by assigning load to a server from the first and second server groups after time T because Dias et al. discloses time intervals of distribution for groups and Devarakonda further incorporates distribution time for servers within the group therefore it is obvious to continually balance the load servers must continually receive request and the server group with servers with less load can receive the greatest amount of request over time thus the

group with the lowest load starts to receive request and groups with greater loads began reception at a later time to better balance the load. At the same time the load will become unbalanced if the same server within the group continues to receive request thus as a server from a new group begins to receive request, a new server is selected from previous groups to receive request as well.

Regarding claim 2, Yu teaches in a communication system network having a plurality of servers (Fig. 1, reference 54 of Yu), each of said plurality of servers having a load level based on serving a number of clients in said communication system network (column 1, lines 9-11 and 28-30 of Yu), a method comprising the steps of: grouping said plurality of servers into a server group G0 (column 4, lines 52-53 of Yu), wherein server groups G0 through G2 respectively have load levels progressively from a least amount of lead level to a most amount of load level (column 3, lines 53-55 and column 6, lines 55-58of Yu); calculating a time period T1 (column 3, line 1 of Yu); assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1 (column 2, lines 25-27 and 58-60 of Yu).

Yu does not teach grouping said plurality of servers into a plurality of server groups G0 through G2. Dias et al. teaches grouping said plurality of servers into a plurality of server groups G0 through G2 (column 2, lines 58-65 and column 3, lines 3-5 and column 4, lines 20-22). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu by grouping said plurality of servers into a plurality of server groups G0 through G2 because the creation of server groups based on server load allows request to be sent to a

group where all servers within the group are just as capable of processing the request without overload thus creating request response time more efficient and maintaining load balance.

Yu in view of Dias et al. does not teach of multiple time periods or how to assign load to a server during time periods other than the initial time period. Devarakonda et al. teaches calculating a time period T1 and T2, wherein said time period T2 is longer than said time period T1; assigning load to a server selected from servers in said server groups G0 and G1 after expiration of said time period T1; assigning load to a server selected from servers in said server groups G0, G1 and G2 after expiration of said time period T2 (column 2, lines 53-57). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu in view of Dias et al. by having multiple time periods and assigning load to a server from each server group during each time period after the initial time period because Dias et al. discloses time intervals of distribution for groups and Devarakonda further incorporates distribution time for servers within the group therefore it is obvious to continually balance the load servers must continually receive request and the server group with servers with less load can receive the greatest amount of request over time thus the group with the lowest load starts to receive request and groups with greater loads began reception at a later time to better balance the load. At the same time the load will become unbalanced if the same server within the group continues to receive request thus as a server from a new group begins to receive request, a new server is selected from previous groups to receive request as well.

Referring to claim 3, Yu teaches in a communication system network having a plurality of servers (Fig. 1, reference 54 of Yu), each of said plurality of servers having a load level based

on serving a number of clients in said communication system network (column 1, lines 9-11 and 28-30 of Yu), a method comprising the steps of: grouping said plurality of servers into a server group G0 (column 4, lines 52-53 of Yu), wherein said server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level (column 3, lines 53-55 and column 6, lines 55-58 of Yu); calculating a time period T1 (column 3, line 1 of Yu); assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1 (column 2, lines 25-27 and 58-60 of Yu).

Yu does not teach grouping said plurality of servers into a plurality of server groups G0 through Gk. Dias et al. teaches grouping said plurality of servers into a plurality of server groups G0 through Gk (column 2, lines 58-65 and column 3, lines 3-5 and column 4, lines 20-22). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu by grouping said plurality of servers into a plurality of server groups G0 through Gk because the creation of server groups based on server load allows request to be sent to a group where all servers within the group are just as capable of processing the request without overload thus creating request response time more efficient and maintaining load balance.

Yu in view of Dias et al does not teach of multiple time periods or how to assign load to a server during time periods other than the initial time period. Devarakonda et al. teaches calculating a plurality of time periods T1 through Tk; assigning load, after expiration of each of said time periods T1 through Tk measured form said initial time, to a server selected from servers in the server groups from G0 and at least one other group selected from said server groups G1 through Gk (column 2, lines 53-57). Therefore, it would have been obvious to one

having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu in view of Dias et al by having multiple time periods and assigning load to a server from each server group during each time period after the initial time period because Dias et al. discloses time intervals of distribution for groups and Devarakonda further incorporates distribution time for servers within the group therefore it is obvious to continually balance the load servers must continually receive request and the server group with servers with less load can receive the greatest amount of request over time thus the group with the lowest load starts to receive request and groups with greater loads began reception at a later time to better balance the load. At the same time the load will become unbalanced if the same server within the group continues to receive request thus as a server from a new group begins to receive request, a new server is selected from previous groups to receive request as well.

Regarding claim 4, Yu teaches in a communication system network having a plurality of servers (Fig. 1, reference 54 of Yu), each of said plurality of servers having a load level based on serving a number of clients in said communication system network (column 1, lines 9-11 and 28-30of Yu), a method comprising the steps of: grouping said plurality of servers into a server group G0 (column 4, lines 52-53 of Yu), wherein said server groups G0 through Gk respectively have load levels progressively from a least amount of load level to a most amount of load level (column 3, lines 53-55 and column 6, lines 55-58of Yu); calculating a time period T1 corresponding to said server group G1 (column 3, line 1 of Yu); assigning load to a server selected from servers in said server group G0 from an initial time until expiration of said time period T1 (column 2, lines 25-27 and 58-60 of Yu).

Yu does not teach grouping said plurality of servers into a plurality of server groups G0 through Gk. Dias et al. teaches grouping said plurality of servers into a plurality of server groups G0 through Gk (column 2, lines 58-65 and column 3, lines 3-5 and column 4, lines 20-22). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu by grouping said plurality of servers into a plurality of server groups G0 through Gk because the creation of server groups based on server load allows request to be sent to a group where all servers within the group are just as capable of processing the request without overload thus creating request response time more efficient and maintaining load balance.

Yu in view of Dias et al. does not teach of multiple time periods corresponding to multiple server groups nor how to assign load to a server during time periods other than the initial time period. Devarakonda et al. teaches calculating a plurality of time periods T1 through Tk corresponding to said server groups G1 through Gk; assigning load, after expiration of each of said time periods T1 through Tk measured from said initial time, to a server selected from servers in a combination of servers including said server group G0 and at least one other server group, in said server groups G1 through Gk, corresponding to an expiring time period (column 2, lines 53-57). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu in view of Dias et al. by having multiple time periods that correspond to the server groups and assigning load to a server selected from servers in a combination of servers including said server group G0 and at least one other server group during each time period after the initial time period because Dias et al. discloses time intervals of distribution for

groups and Devarakonda further incorporates distribution time for servers within the group therefore it is obvious to continually balance the load servers must continually receive request and the server group with servers with less load can receive the greatest amount of request over time thus the group with the lowest load starts to receive request and groups with greater loads began reception at a later time to better balance the load. At the same time the load will become unbalanced if the same server within the group continues to receive request thus as a server from a new group begins to receive request, a new server is selected from previous groups to receive request as well.

Regarding claim 7, Yu teaches the method as recited in claim 4 wherein said grouping of said server groups G0 is based on similarity of load levels among said plurality of servers (column 3, lines 53-55 and column 4, lines 52-53 and column 5, lines 15-17 of Yu).

Yu does not teach said grouping of said plurality of server groups G0 through Gk. Dias et al. teaches said grouping of said plurality of server groups G0 through Gk (column 2, lines 58-65 and column 3, lines 3-5 and column 4, lines 20-22). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu by said grouping of said plurality of server groups G0 through Gk because there would be a group of servers able to handle various load levels.

Regarding claim 10, Yu teaches the method as recited in claim 4 wherein each of said plurality of time periods T1 through Tk is based on load level of server group, a request arrival rate and a server service rate (column 2, lines 12-14 and column 3, lines 18-21 and lines 34-36 of Yu).

Yu does not teach load levels of at least two server groups selected from said plurality of server groups G0 through Gk. Dias et al. teaches load levels of at least two server groups selected from said plurality of server groups G0 through Gk (column 2, lines 58-65 and column 3, lines 3-5 and column 4, lines 20-22). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu by load levels of at least two server groups selected from said plurality of server groups G0 through Gk because there would be a group of servers able to handle various load levels.

Referring to claim 8, Yu teaches the method as recited in claim 4 wherein at least one load assignment in said assigning load to a server in said server group G0 and said assigning load to a server selected from servers in said combination is performed according to a round robin selection method (column 2, line 25-26 of Yu).

Referring to claims 11 and 12, Yu teaches the method as recited in claim 10 wherein said request arrival rate is substituted for an average request arrival rate [of a combination of servers] of said plurality of servers (column 6, line 56 and column 7, lines 7-8 of Yu).

Regarding claims 13 and 14, Yu teaches the method as recited in claim 10 wherein said server service rate is substituted for an average service rate [of a combination of servers] of said plurality of servers (column 5, lines 55-56 and column 5, lines 66-67 through column 6, line 1 of Yu).

4. Claim 5 is rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,078,943 to Yu in view of Dias et al. and further in view of Devarakonda et al. as applied to claims 1-4, 7, 8, and 10-14 above, and further in view of Adelman et al.

Regarding claim 5, Yu in view of Dias et al. and further in view of Devarakonda et al. teaches of server groups (column 4, lines 52-53 of Yu) and time periods (column 3, line 1 of Yu).

Yu in view of Dias et al. and further in view of Devarakonda et al. does not teach that the plurality of time periods T1 through Tk each are based on a difference between load levels of at least two server groups. Adelman et al. teaches the method as recited in claim 4 wherein said plurality of time periods T1 through Tk each is based on a difference between load levels of at least two server groups in said plurality of server groups G0 through Gk (column 8, lines 36-40). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further the method and apparatus for dynamic interval-based load balancing of Yu in view of Dias et al. and further in view of Devarakonda et al. by having the time periods be based on the difference between load levels of the server groups because the time periods limit the amount of time to assign load to the servers of multiple server groups therefore the time period needs to be longer for servers (from the higher end server groups) being assigned a greater load and as the higher end server groups are included the load difference will increase which will increase the time period.

5. Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,078,943 to Yu in view of Dias et al. and further in view of Devarakonda et al. as applied to claims 1-4, 7, 8, and 10-14 above, and further in view of Attanasio et al. and Fine.

Referring to claim 6, Yu in view of Dias et al. and further in view of Devarakonda et al. teaches of assigning loads to servers within a server group for a designated time period (column 1, lines 11-12 and column 4, lines 52-53 and column 5, line 1-3 of Yu).

Yu in view of Dias et al. and further in view of Devarakonda et al. does not teach of receiving an update of the load level of the servers and regrouping the servers, and reassigning the load based on this update. Attanasio et al. teaches the method as recited in claim 4 further comprising the step of: receiving an update of load level of at least one of said plurality of servers in said plurality of server groups G0 through Gk (column 3, lines 46-48 of Attanasio et al.); repeating said grouping to produce a new plurality of server groups G0 through Gk based on said update of load level (column 3, lines 60-62 of Attanasio et al.); and assigning load to a server selected from servers in said new server group G0 from said reset initial time until expiration of said new time period T1; assigning load, after expiration of each of said new time periods T1 through Tk measured from said reset initial time, to a server selected from servers in a combination of servers including said new server group G0 and at least one other server, in said new server groups G1 through Gk, corresponding to an expiring time period (column 8, lines 25-27 of Attanasio et al.). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu in view of Dias et al. and further in view of Devarakonda et al. by receiving an update of the load level of the servers and regrouping the servers, and reassigning the load based on this update because with this system there is a chance of an imbalance of the load if servers within a server group are assigned a load greater than the load threshold for that server group therefore a server load update and the repetition of grouping servers and assigning load is needed to keep the load balanced within the system.

Yu in view of Dias et al. and further in view of Devarakonda et al. and further in view of Attanasio et al. does not teach recalculating the time periods and resetting the initial time. Fine

teaches repeating said calculating said plurality of time periods to produce a new plurality of time periods T1 through Tk corresponding to said new plurality of server groups G0 through Gk; resetting said initial time to a reset initial time (column 10, lines 1-2 and column 16, lines 25-27 of Fine). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load balancing of Yu in view of Dias et al. and further in view of Devarakonda et al. by recalculating the time periods and resetting the initial time because with repetition of grouping servers and assigning load a new set of time periods are needed because the time period is related to the server group.

1. Claim 9 is rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,078,943 to Yu in view of Dias et al. and further in view of Devarakonda et al. as applied to claims 1-4, 7, 8, and 10-14 above, and further in view of Applicant Admitted Prior Art (AAPA).

Regarding claim 9, Yu in view of Dias et al. and further in view of Devarakonda et al. teaches assigning load to a server in said server group (column 1, lines 11-12 and column 4, lines 52-53 of Yu).

Yu in view of Dias et al. and further in view of Devarakonda et al. does not teach of load assignment according to a random selection method. AAPA teaches the method as recited in claim 4 wherein at least one load assignment in said assigning load to a server in said server group G0 and said assigning load to a server selected from servers in said combination is performed according to a random selection method (pg.4, lines 8-10 and pg. 11, lines 10-11). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to further modify the method and apparatus for dynamic interval-based load

balancing of Yu in view of Dias et al. and further in view of Devarakonda et al. by assigning load to the servers according to a random selection method because this helps to avoid the server groups moving together as in the round robin method.

Conclusion

2. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. The following patent is cited to further show the state of the art with respect to load balancing in general:

US Pat No. 6,324,177 to Howes et al.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to April L Baugh whose telephone number is 703-305-5317. The examiner can normally be reached on Monday-Friday 8:30am-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Rupal D Dharia can be reached on 703-305-4003. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is 703-305-3900.

ALB

RUPAL DHARIA
SUPERVISORY PATENT EXAMINER